EURASIP Journal on Wireless Communications and Networking 2004:2, 322–334

Blind Channel Estimation for Space-Time Coded WCDMA

Youngchul Sung
School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853, USA
Email: ys87@ece.cornell.edu

Lang Tong
School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853, USA
Email: ltong@ece.cornell.edu

Ananthram Swami
Army Research Laboratory, 2800 Powder Mill Read, Adelphi, MD 20783, USA
Email: a.swami@ieee.org

Received 29 November 2003; Revised 20 April 2004

A new blind channel estimation technique is proposed for space-time coded wideband CDMA systems using aperiodic and possibly multirate spreading codes. Using a decorrelating front end, the received signal is projected onto a subspace from which channel parameters can be estimated up to a rotational ambiguity. Exploiting the subspace structure of the WCDMA signaling and the orthogonality of the unitary space-time codes, the proposed algorithm provides a blind channel estimate via least squares. A new identifiability condition is established under the assumption that the system is not heavily loaded. The mean square error of the estimated channel is compared with the Cramér-Rao bound, and the bit error rate (BER) performance of the proposed algorithm is compared with that of differential schemes.

Keywords and phrases: space-time coding, long code CDMA, least squares, blind channel estimation.

1. INTRODUCTION

Future wireless systems will require high rate transmission of multimedia data over time-varying fading channels. This is especially the case for the downlink where a mix of voice, low rate data, and possibly images are transmitted to mobile users. To increase the capacity and provide reliable communication over fading channel, diversity techniques in space and time are expected to play a crucial role [1, 2, 3, 4]. A variety of space-time coding schemes have been proposed with multiple transmit antennas and a single or multiple receive antennas (e.g., [5, 6, 7]). Indeed, the 3G wireless standards support base station transmit diversity at the WCDMA physical layer.

Many space-time techniques, the popular Alamouti scheme in particular, are designed for coherent detection where channel estimation is necessary. There is a substantial literature, for example, [8, 9, 10], addressing the channel estimation issue for (space-time coded) multiple-input multiple-output (MIMO) systems, ranging from standard training-based techniques that rely on pilot symbols in the data stream to blind and semiblind methods where observations corresponding to data and pilots (if they exist) are used jointly. Noncoherent detection schemes for space-time coded systems have also been proposed based on differential or sequential decoding [11, 12, 13]. These methods avoid the need for channel estimation by introducing structure in the transmitted symbol stream. The receiver can demodulate the transmitted symbols directly by exploiting the embedded structure. Although these methods increase bandwidth efficiency by eliminating the necessity for training symbols, and are robust to fast fading, they suffer from performance degradation due to the error propagation problem.

For WCDMA systems, several spatial diversity schemes such as *orthogonal transmit diversity* (OTD) [14], *space-time spreading* (STS) [15], and *space-time block coding based transmit diversity* (STTD) have been proposed and adopted. These diversity techniques provide additional reliability on top of the robustness of CDMA systems against multiuser interference. In this paper, we focus on WCDMA systems with space-time block coding based transmit diversity. The challenge of

channel estimation in such a wideband system is twofold. First, the WCDMA is a multirate system where the delay spread may exceed several symbol intervals causing severe multipath fading and intersymbol interference; the channel is a MIMO system with memory. Second, the increase in the number of channel parameters, due to the use of multiple antennas, makes the conventional training-based scheme less reliable and more prone to multiaccess interference. Fortunately, WCDMA also offers signal structures that could be exploited in an estimation scheme.

Blind estimation or detection algorithms have been proposed for space-time coded CDMA systems. For example, a blind channel estimation technique based on the Capon receiver or the minimum output variance technique for flat fading channels, with two spreading codes per user, was proposed in [16]. In this paper, we propose a blind channel estimation technique for frequency-selective fading channels, with a single spreading code per user. The proposed method requires no more than two pilot symbols per user per slot. (This is the same number of pilot symbols as in differential detection schemes.) The proposed algorithm exploits the subspace structure of the long code WCDMA transmission and the orthogonality of the unitary codes, for example, the Alamouti code. As a subspace technique, the proposed algorithm is based on the front-end processing, and requires the code matrix to be invertible in the case of the decorrelating front ends. The proposed method can obtain channel estimates quickly using only one slot, which allows us to deal with rapidly fading channels. Using a rake structure, our technique is compatible with the standard receiver front ends that suppress multiaccess interference, and perform decoding for each user separately.

The paper is organized as follows. The data model of a space-time coded long code CDMA system is described in Section 2. In Section 3, the new blind channel estimation method is proposed based on decorrelation and an identifiability condition is established. Several extensions are also discussed. In Section 4, detection schemes are briefly discussed. In Section 5, the performance of the proposed method is compared with the Cramér-Rao Bound (CRB) through Monte Carlo simulations and the bit error rate (BER) of the proposed method is compared with that of differential detection schemes.

1.1. Notation

The notations are standard. Vectors and matrices are written in boldface with matrices in capitals. We reserve $\mathbf{I}_m$ for the identity matrix of size m (the subscript is included only when necessary). For a random vector $\mathbf{x}$, $\mathbb{E}(\mathbf{x})$ is the mathematical expectation of $\mathbf{x}$. The notation $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ means that $\mathbf{x}$ is (complex) Gaussian with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. For a complex quantity α, α^* and $\mathrm{Re}(\alpha)$ denote the complex conjugate and the real part of α, respectively. Operations $(\cdot)^{\mathrm{T}}$ and $(\cdot)^H$ indicate transpose and Hermitian transpose, respectively. $\mathrm{tr}(\cdot)$ denotes the trace of a matrix. $\mathrm{diag}(\mathbf{X}_1, \ldots, \mathbf{X}_N)$ is a block diagonal matrix with $\mathbf{X}_1, \ldots, \mathbf{X}_N$ as its diagonal blocks. Given a matrix $\mathbf{X}$, $\mathbf{X}^\dagger$ is the Moore-Penrose pseudoinverse and $\mathbf{X} \otimes \mathbf{Y}$ is the Kronecker product of $\mathbf{X}$ and $\mathbf{Y}$. For a matrix (vector) $\mathbf{X}$, we use $\|\mathbf{X}\|$ for the 2-norm and $\|\mathbf{X}\|_F$ for the Frobenius norm.

2. DATA MODEL

We consider STTD that requires only a single spreading code for each user. Specifically, we consider a WCDMA system with the Alamouti coding scheme [5]. We assume two transmit antennas and a single receive antenna, K asynchronous users with aperiodic spreading codes, and slotted transmissions.

At the transmitter, user i transmits two data sequences $\{s_{im}^{(1)}\}_{m=1}^{M_i}$ and $\{s_{im}^{(2)}\}_{m=1}^{M_i}$, one through each antenna, in each slot. The data sequence for user i is space-time encoded as

$$\begin{aligned} s_{im}^{(1)} &= s_{im}, \\ s_{i,m+1}^{(1)} &= s_{i,m+1}, \\ s_{im}^{(2)} &= -s_{i,m+1}^{(1)*}, \\ s_{i,m+1}^{(2)} &= s_{im}^{(1)*}, \quad m = 1, 3, \ldots, M_i - 1, \end{aligned} \tag{1}$$

where $s_{im} \triangleq s_i(mT_i)$ is the input data sequence, $s_{im}^{(j)} \triangleq s_i^{(j)}(mT_i)$, $j = 1, 2$, the encoded data sequence for transmit antenna j, T_i the symbol interval, and M_i the slot size for user i. Each data sequence is spread by a user-specific long spreading code $c_i(t)$ with spreading gain G_i, followed by a chip rate pulse-shaping filter, and transmitted through the corresponding antenna. Note that the data sequences for the two transmit antennas are spread by the same spreading code here. The separation of the two antenna signals is possible with a single spreading code due to the space-time encoding.[1]

We assume that the channel for each transmit-receive pair of each user does not change for a single slot period, and model it by a complex finite impulse response (FIR) filter with taps separated by multiples of the chip interval. The continuous-time channel impulse response of the path from transmitter j to the single receiver for user i is given by

$$h_i^{(j)}(\tau) = \sum_{l=1}^{L_i^{(j)}} h_{il}^{(j)} \delta\left(\tau - lT_c - d_i^{(j)} T_c\right), \tag{2}$$

where $h_{il}^{(j)}$ is the lth path gain for transmit-receive pair j for user i and $T_c = T_i/G_i$ is the chip interval. We assume that the channel order $L_i^{(j)}$ and the delay $d_i^{(j)}$ from the slot reference are known. We set L_i as the maximum of $\{L_i^{(j)}\}_{j=1,2}$ and d_i as the minimum of $\{d_i^{(j)}\}_{j=1,2}$. When the channel is sparse, it is more efficient to model the channel as separate

[1] When a different spreading code is used for each antenna, this can be considered just as two different CDMA users and the space-time coding is not necessary to achieve the spatial diversity due to the separation capability of the spreading codes. However, this method requires twice as many spreading codes as the system considered here.

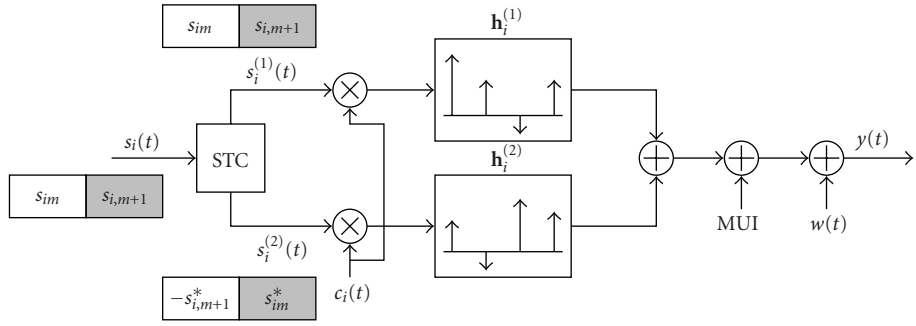


Figure 1: CDMA system with space-time coding using two transmit antennas (STC: space-time encoder).

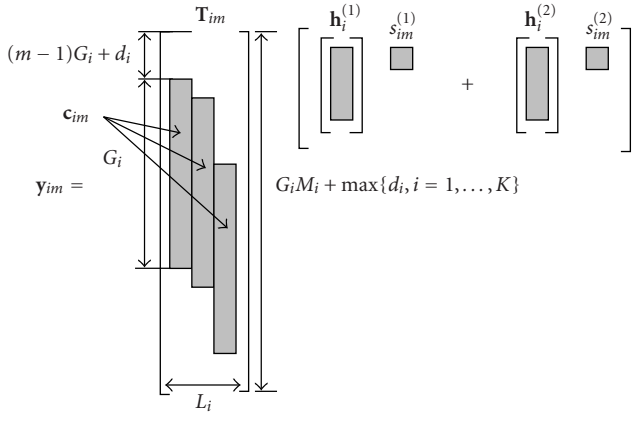


Figure 2: Noiseless single symbol output $\mathbf{y}_{im}$.

clusters of multipaths. In that case, we assume that the approximate locations of these clusters are known. We assume that the transmitted signal is also corrupted by other user interference and additive noise in the channel. The overall system model is described in Figure 1.

At the receiver, we let $y(t)$ pass through the chip-matched filter, and sample it at the chip rate. Stacking the chip rate samples, we obtain the discrete-time received signal vector. First, we consider $\mathbf{y}_{im}$ that corresponds to the noiseless output due to the mth symbol of user i. $\mathbf{y}_{im}$ is given by

$$\mathbf{y}_{im} = \mathbf{T}_{im}\left[\mathbf{h}_i^{(1)}s_{im}^{(1)} + \mathbf{h}_i^{(2)}s_{im}^{(2)}\right], \tag{3}$$

where $\mathbf{h}_i^{(j)} \triangleq [h_{i1}^{(j)},\ldots,h_{iL_i}^{(j)}]^{\mathrm{T}}$ is the vector containing all multipath coefficients of antenna pair j and $\mathbf{T}_{im}$ is the Toeplitz matrix whose first column is made of $(m-1)G_i+d_i$ zeros followed by the code vector $\mathbf{c}_{im}$ (the mth segment of G_i chips of the spreading code of user i) and additional zeros that make the size of $\mathbf{y}_{im}$ the total number of chips of the entire slot plus $\max\{d_i,\ i = 1,\ldots,K\}$ (see Figure 2). Here, we assume that the slot size is fixed for different spreading gains, that is, $G_1M_1 = \cdots = G_KM_K$.

Since the channel is linear, the total received noiseless signal for user i is given by the sum of $\mathbf{y}_{im}$, $m = 1,\ldots,M_i$, as

$$\begin{aligned}\mathbf{y}_i &= \sum_{m=1}^{M_i}\mathbf{T}_{im}\left[\mathbf{h}_i^{(1)}s_{im}^{(1)} + \mathbf{h}_i^{(2)}s_{im}^{(2)}\right]\\ &= \mathbf{T}_i\left(\mathbf{I}_{M_i}\otimes\left[\mathbf{h}_i^{(1)}\ \ \mathbf{h}_i^{(2)}\right]\right)\mathbf{s}_i,\\ \mathbf{s}_i &\triangleq \left[s_{i1}^{(1)}, s_{i1}^{(2)}, s_{i2}^{(1)}, s_{i2}^{(2)},\ldots, s_{iM_i}^{(2)}\right]^{\mathrm{T}},\\ \mathbf{T}_i &\triangleq \left[\mathbf{T}_{i1},\mathbf{T}_{i2},\ldots,\mathbf{T}_{iM_i}\right],\end{aligned} \tag{4}$$

where $\mathbf{T}_i$ is the code matrix of user i and has a special block shifting structure. Including all users and noise, we have the complete matrix model given by

$$\begin{aligned}\mathbf{y} &= \left[\mathbf{T}_1\cdots\mathbf{T}_K\right]\mathrm{diag}\left(\mathbf{I}_{M_1}\otimes\mathbf{H}_1,\ldots,\mathbf{I}_{M_K}\otimes\mathbf{H}_K\right)\mathbf{s} + \mathbf{w}\\ &= \mathbf{T}\mathcal{D}(\mathbf{H})\mathbf{s} + \mathbf{w},\end{aligned} \tag{5}$$

where the overall code matrix $\mathbf{T}$ of size $(G_1M_1 + \max\{d_i\}) \times \sum_{i=1}^{K} M_iL_i$ is composed of the code matrices of all K users, $\mathbf{s}$ includes all symbols of both transmitters for all users, and

$$\mathbf{H}_i \triangleq \left[\mathbf{h}_i^{(1)}\ \ \mathbf{h}_i^{(2)}\right]. \tag{6}$$

$\mathbf{H}_i$ contains the channels of both transmit-receive pairs for user i. The matrix $\mathcal{D}(\mathbf{H})$ is block diagonal with $\mathbf{I}_{M_i}\otimes\mathbf{H}_i$ as the block element. (See Figure 3 for the example of two-user case.) The additive noise is denoted by $\mathbf{w}$.

We will make the following assumptions.

(A1) The code matrix $\mathbf{T}$ is known.

(A1′) The code matrix $\mathbf{T}$ has full column rank.

(A2) The channel matrix $\mathbf{H}_i$ is full column rank.

(A3) The noise vector is complex Gaussian $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$ with possibly unknown variance σ^2.

Assumption (A1) implies that the receiver knows the codes for all users as well as the delay d_i and the maximum channel order L_i. Rough knowledge of the delay d_i is enough since we can overparameterize the channel to accommodate the delay uncertainty. When the knowledge of other users' codes is not available, we model other user interference as Gaussian noise.

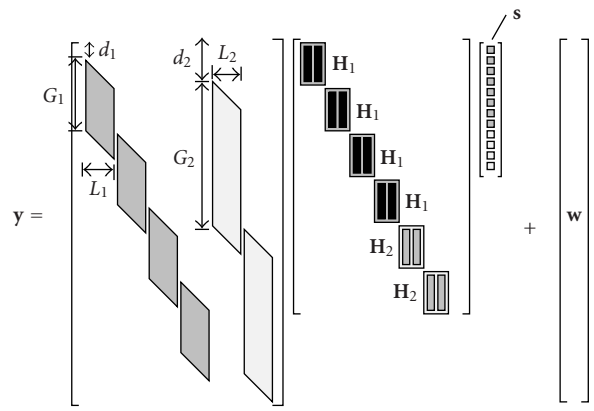


Figure 3: Multiuser matrix model for the received signal.

For the downlink case, the relative delay d_i and the number of multipaths L_i are the same for all users. Since the downlink spreading usually uses orthogonal codes and the orthogonality between signals of different users is disturbed only by multipaths, other user interference is not severe after equalizing the multipath effect. For the case of multiple spreading codes for a single user, we can model all the codes in the code matrix. Assumption (A1′) is sufficient but not necessary for the channel to be identifiable. Assumption (A2) requires that the number of multipaths be at least two (this is reasonable for typical wireless channels) and the two transmit-receive pairs have uncorrelated channels. The latter condition is usually guaranteed for well-designed spatial diversity systems by proper antenna spacing.

3. BLIND CHANNEL ESTIMATION

In this section, we propose a blind channel estimation that identifies the channel for both antenna pairs simultaneously up to unitary rotational ambiguity with one slot observation. The method is based on the decorrelation of user signals that projects the received signal onto a subspace from which the channels of both transmit-receive pairs are estimated using a low-rank decomposition. Blind estimation is possible due to the unitary property of the space-time codes. The proposed method combines two consecutive symbols, and eliminates the unknown symbols by exploiting this unitary property. We assume that the channel and symbols are deterministic parameters.

3.1. Blind algorithm

3.1.1. Front-end processing

We consider decorrelator, conventional matched filter, and regularized decorrelator as the front end. The decorrelator is basically assumed for the algorithm construction. However, other front ends can be applied to the same algorithm depending on the situation and their performances are also evaluated in Section 5. The decorrelating front-end $\mathbf{T}^\dagger$ can be efficiently implemented using a state-space inversion technique that significantly reduces the complexity and storage requirement by exploiting the structure of the code matrix [17].

The output of the decorrelator is given in vector form by

$$\begin{aligned}\mathbf{z} &= \mathbf{T}^\dagger \mathbf{y} = \mathcal{D}(\mathbf{H})\mathbf{s} + \mathbf{n} \\ &= \operatorname{diag}\left(\mathbf{I}_{M_1} \otimes \mathbf{H}_1, \ldots, \mathbf{I}_{M_K} \otimes \mathbf{H}_K\right)\mathbf{s} + \mathbf{n},\end{aligned} \tag{7}$$

where $\mathbf{n} = \mathbf{T}^\dagger \mathbf{w}$ is now colored. We segment $\mathbf{z}$ and obtain subvector $\mathbf{z}_{im}$ of size L_i, $m = 1, 2, \ldots, M_i$. In the case of equal spreading gain and equal channel order ($M_1 = \cdots = M_K = M$ and $L_1 = \cdots = L_K = L$), $\mathbf{z}_{im}$ is the $((i-1)M + m)$th L-dimensional subvector of $\mathbf{z}$. The subvectors corresponding to two consecutive symbols $2n-1$, $2n$ of user i are given by

$$\begin{aligned}\mathbf{z}_{i,2n-1} &= \mathbf{H}_i \begin{bmatrix} s_{i,2n-1} \\ -s_{i,2n}^* \end{bmatrix} + \mathbf{n}_{i,2n-1}, \\ \mathbf{z}_{i,2n} &= \mathbf{H}_i \begin{bmatrix} s_{i,2n} \\ s_{i,2n-1}^* \end{bmatrix} + \mathbf{n}_{i,2n},\end{aligned} \tag{8}$$

where $n = 1, 2, \ldots, M_i/2$ (see Figure 3). Rewriting the two vectors in a matrix form yields

$$\mathbf{Z}_{in} \triangleq \begin{bmatrix} \mathbf{z}_{i,2n-1} & \mathbf{z}_{i,2n} \end{bmatrix} = \mathbf{H}_i \mathbf{S}_{in} + \mathbf{N}_{in}, \tag{9}$$

where $\mathbf{H}_i$ contains the unknown channel vector for each transmit-receive pair as described in (6), $\mathbf{N}_n = \begin{bmatrix} \mathbf{n}_{i,2n-1} & \mathbf{n}_{i,2n} \end{bmatrix}$, and

$$\mathbf{S}_{in} = \begin{bmatrix} s_{i,2n-1} & s_{i,2n} \\ -s_{i,2n}^* & s_{i,2n-1}^* \end{bmatrix}. \tag{10}$$

Here, $\mathbf{S}_{in}$ belongs to the space-time code $\mathcal{S}$. Notice that the rearranged front-end output (9) in the CDMA with multipaths has an equivalent signal structure for (nonspread) MIMO channel for 2 transmit antennas and L_i receive antennas with flat fading for each transmit-receive pair.

3.1.2. Low-rank decomposition

We utilize the orthogonal property of unitary space-time codes including the Alamouti scheme to eliminate the unknown symbols. Due to the unitary property of the codes, we have

$$\mathbf{S}_{in}\mathbf{S}_{in}^H = \mathbf{S}_{in}^H \mathbf{S}_{in} = \alpha_{in}\mathbf{I}, \tag{11}$$

where $\alpha_{in} = |s_{i,2n-1}|^2 + |s_{i,2n}|^2$. For the case of symbols with constant energy, α_{in} is fixed for all n and known beforehand.

In noiseless case, it is easily seen that multiplying $\mathbf{Z}_{in}$ by its Hermitian eliminates the unknown symbols to make blind identification possible. In noisy case, utilizing all the observations, we can form a least squares estimate of the channel matrix. Let $\mathbf{Z}_i \triangleq [\mathbf{Z}_{i1}, \mathbf{Z}_{i2}, \ldots, \mathbf{Z}_{i,M_i/2}]$. Then, we have

$$\mathbf{Z}_i = \mathbf{H}_i \mathbf{S}_i + \mathbf{N}_i, \tag{12}$$

where

$$\begin{aligned}\mathbf{S}_i &\triangleq [\mathbf{S}_{i1}, \mathbf{S}_{i2}, \ldots, \mathbf{S}_{i,M_i/2}], \\ \mathbf{N}_i &\triangleq [\mathbf{N}_{i1}, \mathbf{N}_{i2}, \ldots, \mathbf{N}_{i,M_i/2}].\end{aligned} \tag{13}$$

The least squares estimator for $\mathbf{H}_i$ and $\mathbf{S}_i$ is given by

$$[\widehat{\mathbf{H}_i}, \widehat{\mathbf{S}}_{in}] = \underset{\mathbf{H}_i, \{\mathbf{S}_{in} \in \mathcal{S}\}}{\arg\min} ||\mathbf{Z}_i - \mathbf{H}_i \mathbf{S}_i||_F^2. \tag{14}$$

Since the exact solution of (14) is not tractable in a closed form [8], we apply a suboptimal two-step approach: we first estimate the channel only, and then detect the symbols using the estimated channel. (See Section 4 for the subsequent symbol detection.) Solving (14) by relaxing the constraint of $\mathbf{S}_{in}$ on the signal constellation, the subspace of $\mathbf{H}_i$ is obtained. Notice that $\mathbf{H}_i\mathbf{S}_i$ is rank-deficient if $L_i > 2$ since $\mathbf{H}_i$ has rank two by its construction.[2] Hence, the subspace of $\mathbf{H}_i$ is obtained by low rank approximation via singular value decomposition (SVD) of $\mathbf{Z}_i$ [18]. Let the SVD of $\mathbf{Z}_i$ be given by

$$\mathbf{Z}_i = \mathbf{U}_i \Sigma_i \mathbf{V}_i^H. \tag{15}$$

Then, the estimate for the product of channel and symbol is given by

$$\widehat{\mathbf{H}_i \mathbf{S}_i} = \sum_{j=1}^{2} \sigma_{ij} \mathbf{u}_{ij} \mathbf{v}_{ij}^H, \tag{16}$$

where σ_{ij} is the singular values in Σ_i, and $\mathbf{u}_{ij}$ and $\mathbf{v}_{ij}$ are the jth column of $\mathbf{U}_i$ and $\mathbf{V}_i$, respectively. Now, we utilize the orthogonality (11) of the space-time code and eliminate $\mathbf{S}_i$ from (16). Since $\mathbf{S}_i\mathbf{S}_i^H = (\sum_{n=1}^{M_i/2} \alpha_{in})\mathbf{I}$, multiplying the estimate for the product by its Hermitian gives

$$\begin{aligned} \alpha_i \widehat{\mathbf{H}_i \mathbf{H}_i^H} &= \sum_{j=1}^{2} \sigma_{ij}^2 \mathbf{u}_{ij} \mathbf{u}_{ij}^H \\ &= \widetilde{\mathbf{U}}_i \begin{bmatrix} \sigma_{i1}^2 & 0 \\ 0 & \sigma_{i2}^2 \end{bmatrix} \widetilde{\mathbf{U}}_i^H, \end{aligned} \tag{17}$$

where $\alpha_i = \sum_{n=1}^{M_i/2} \alpha_{in}$ and $\widetilde{\mathbf{U}}_i = [\mathbf{u}_{i1}, \mathbf{u}_{i2}]$. Finally, the estimate for $\mathbf{H}_i$ is given by

$$\hat{\mathbf{H}}_i = \frac{1}{\sqrt{\alpha_i}} \widetilde{\mathbf{U}}_i \tilde{\Sigma}_i \mathbf{Q}_i, \tag{18}$$

where $\tilde{\Sigma}_i = \mathrm{diag}(\sigma_{i1}, \sigma_{i2})$ and $\mathbf{Q}_i$ is an unknown 2×2 unitary matrix. The rotational ambiguity in the above estimate must be removed by either incorporating prior knowledge of the symbol or by using pilot symbols. The singular values and left singular vectors of $\mathbf{Z}_i$ can be obtained using a smaller matrix $\mathbf{R}_i$ defined as

$$\mathbf{R}_i \triangleq \sum_{n=1}^{M_i/2} \mathbf{Z}_{in} \mathbf{Z}_{in}^H, \tag{19}$$

where its SVD is given by

$$\mathbf{R}_i = \mathbf{U}_i \Sigma_i^2 \mathbf{U}_i^H. \tag{20}$$

[2] $L_i \geq 2$ is sufficient for the algorithm.

3.2. Identifiability

We have so far assumed that the overall code matrix $\mathbf{T}$ has full column rank, (A1′), and therefore invertible from the left, that is, $\mathbf{T}^\dagger\mathbf{T} = \mathbf{I}$. This assumption is usually valid for systems with large spreading gains or small delay spreads. (For the case of equal spreading gain and channel order, the size of the code matrix $\mathbf{T}$ is $GM \times LMK$. We need $G \geq LK$). Under this assumption, it is clear that each user's channel is identifiable up to a rotational matrix ambiguity. When the spreading gain is small and the system is heavily loaded, $\mathbf{T}$ can be singular. We present a general identifiability condition for the proposed method that is independent of the channel parameters.

Proposition 1. *Let* $\widetilde{\mathbf{T}}_{in} \triangleq [\mathbf{T}_{i,2n-1} \;\; \mathbf{T}_{i,2n}]$ *be the matrix composed of two consecutive code matrices of user i for symbol* $2n-1$, $2n$, *and* $\check{\mathbf{T}}_{in}$ *the submatrix of* $\mathbf{T}$ *after removing* $\widetilde{\mathbf{T}}_{in}$. *The channel matrix* $\mathbf{H}_i$ *is identifiable up to a rotational ambiguity in the noiseless case if* $\mathbf{T}$ *is a tall matrix and there exists an n such that*

$$\mathcal{C}(\widetilde{\mathbf{T}}_{in}) \bigcap \mathcal{C}(\check{\mathbf{T}}_{in}) = \{\mathbf{0}\}, \tag{21}$$

where $\mathcal{C}(\cdot)$ *denotes the column space of a matrix.*

Proof. If (21) holds for some n, then the range space of $\mathbf{T}$ can be decomposed into the sum of two subspaces, that is, there exists a matrix $\mathbf{V}$ with rank($\mathbf{T}$)–rank($\widetilde{\mathbf{T}}_{in}$) linearly independent columns such that

$$\mathcal{C}([\widetilde{\mathbf{T}}_{in} \;\; \mathbf{V}]) = \mathcal{C}(\mathbf{T}). \tag{22}$$

Let $\mathcal{T} \triangleq [\widetilde{\mathbf{T}}_{in} \;\; \mathbf{V}]$. We have, in the noiseless case,

$$\mathcal{T}^\dagger \mathbf{y} = \begin{bmatrix} * \\ \mathbf{h}_i^{(1)} s_{i,2n-1} - \mathbf{h}_i^{(2)} s_{i,2n}^* \\ \mathbf{h}_i^{(1)} s_{i,2n} + \mathbf{h}_i^{(2)} s_{i,2n-1}^* \\ * \end{bmatrix}. \tag{23}$$

Then, we form $\mathbf{Z}_{in}$ in (9). This implies that $\mathbf{H}_i$ is identifiable up to a rotational ambiguity. □

Since (21) needs to hold only for some n, the use of long codes makes the identifiability condition easy to satisfy. For the downlink case, the condition is easier to satisfy since we have more choices over i.

It is easily seen that any tall code matrix $\mathbf{T}$ has the null space of $\{\mathbf{0}\}$ in the single-user case due to the special block Toeplitz structure. (See Figure 3.) Hence, $\mathbf{T}$ has full column rank and (21) is satisfied in the single-user case. In the multiple-user case, however, it is not easy to have closed-form results on the validity of the condition on $\mathbf{T}$ since it depends on the values of the spreading codes as well as the structure of the matrix. Hence, we checked the validity of the condition through simulation. We evaluated the

condition number of the code matrix $\mathbf{T}$ for random realizations of user spreading codes. The distribution of the condition number as a function of parameters, such as the spreading gain, channel order, and number of users, is shown in Section 5. The simulation shows that for systems with well-designed spreading codes and reasonable load the code matrix is well conditioned and the identifiability condition is satisfied.

3.3. Resolving the rotational ambiguity

The unknown unitary matrix $\mathbf{Q}_i$ in (18) and (30) needs to be resolved for coherent detection of symbols. This can be done using only two consecutive pilot symbols. We formulate a least squares problem for estimating $\mathbf{Q}_i$ using only the observation corresponding to pilot symbols. The estimate for $\mathbf{Q}_i$ is given, from (9) and (18), by

$$\begin{aligned}\hat{\mathbf{Q}}_i &= \arg\min_{\mathbf{Q}\in\mathbb{C}^{2\times 2}} ||\mathbf{Z}_{ip} - \mathbf{H}_i\mathbf{S}_{ip}||_F^2 \\ &= \arg\min_{\mathbf{Q}\in\mathbb{C}^{2\times 2}} \left|\left|\mathbf{Z}_{ip} - \frac{1}{\sqrt{\alpha_i}}\tilde{\mathbf{U}}_i\tilde{\Sigma}_i\mathbf{Q}\mathbf{S}_{ip}\right|\right|_F^2 \\ &= \arg\min_{\mathbf{Q}\in\mathbb{C}^{2\times 2}} \left|\left|\mathbf{Z}_{ip}\mathbf{S}_{ip}^H - \frac{\alpha_{i1}}{\sqrt{\alpha_i}}\tilde{\mathbf{U}}_i\tilde{\Sigma}_i\mathbf{Q}\right|\right|_F^2\end{aligned} \tag{24}$$

under the constraint

$$\mathbf{Q}\mathbf{Q}^H = \mathbf{I}. \tag{25}$$

For the example of two pilot symbols in the beginning of the slot, $\alpha_{i1} = (|s_{i1}|^2 + |s_{i2}|^2)$ and the pilot-related matrices $\mathbf{Z}_{ip}$ and $\mathbf{S}_{ip}$ are given as

$$\mathbf{Z}_{ip} = [\mathbf{z}_{i1}, \mathbf{z}_{i2}], \qquad \mathbf{S}_{ip} = \begin{bmatrix} s_{i1} & s_{i2} \\ -s_{i2}^* & s_{i1} \end{bmatrix}, \tag{26}$$

where s_{i1}, s_{i2} are two pilot symbols for user i.

Proposition 2. *The least squares estimator of* $\mathbf{Q}$ *for* (24) *is given by*

$$\hat{\mathbf{Q}} = \mathbf{U}_Q\mathbf{V}_Q^H, \tag{27}$$

where $\mathbf{U}_Q$ *and* $\mathbf{V}_Q$ *are obtained by SVD of the following matrix, that is,*

$$\frac{\alpha_{i1}}{\sqrt{\alpha_i}}(\tilde{\mathbf{U}}_i\tilde{\Sigma}_i)^H\mathbf{Z}_{ip}\mathbf{S}_{ip}^H = \mathbf{U}_Q\Sigma_Q\mathbf{V}_Q^H. \tag{28}$$

Proof. See the appendix. □

For multiple-pilot symbol blocks, we can formulate the least squares problem to incorporate all the pilot symbols similar to (12).

3.4. Extensions

Since the noise $\mathbf{n}_{im}$ after the decorrelation is colored, a bias is introduced in estimation. We can apply whitening to remove the bias. The expectation of $\mathbf{R}_i$ in (19) is given by

$$\begin{aligned}\mathbb{E}\{\mathbf{R}_i\} &= \alpha_i\mathbf{H}_i\mathbf{H}_i^H + \sigma^2\Delta_i, \\ \Delta_i &= \sum_{m=1}^{M_i}\Sigma_{im},\end{aligned} \tag{29}$$

where Σ_{im} is the diagonal block of $\mathbf{T}^\dagger(\mathbf{T}^\dagger)^H$ with size $L_i \times L_i$ corresponding to the mth symbol of user i. The whitened estimator is given as

$$\hat{\mathbf{H}}_i = \frac{1}{\sqrt{\alpha_i}}\Delta_i^{1/2}\tilde{\Gamma}_i\tilde{\mathbf{S}}_i^{1/2}\mathbf{Q}_i, \tag{30}$$

where $\Delta_i^{1/2}$ is the Cholesky factor of Δ_i, the SVD of the whitened $\mathbf{R}_i$ is given by

$$\Delta_i^{-1/2}\mathbf{R}_i\Delta_i^{-H/2} = \Gamma_i\mathbf{S}_i\Gamma_i^H, \tag{31}$$

and $\tilde{\Gamma}_i$, $\tilde{\mathbf{S}}_i$ are similarly defined as in (18).

For the downlink case, all user signals go through the same channel, that is, $\mathbf{H}_1 = \cdots = \mathbf{H}_K$. We can improve the estimator performance by exploiting this. We combine the matrix $\mathbf{R}_i$ of all users and apply the same subspace decomposition:

$$\begin{aligned}\mathbf{R} &= \frac{1}{K}\sum_{i=1}^{K}\mathbf{R}_i \\ &= \frac{1}{K}\sum_{i=1}^{K}\sum_{n=1}^{M_i/2}\mathbf{Z}_{in}\mathbf{Z}_{in}^H, \\ \Delta &= \frac{1}{K}\sum_{i=1}^{K}\Delta_i.\end{aligned} \tag{32}$$

This process further improves the performance by averaging out the noise as shown in Section 5.

Even if the algorithm is derived using the decorrelator as the front end, we can apply the same subspace technique to different front-ends depending on the situation. For the case of large spreading factors, the proposed method can be applied with the conventional matched filter $\mathbf{T}^H$ without significant performance loss. When the noise level is high, we can use the regularized decorrelator, given by

$$(\mathbf{T}^H\mathbf{T} + \sigma^2\mathbf{I})^{-1}\mathbf{T}^H, \tag{33}$$

to reduce the noise enhancement at the inversion step. As shown in (33), the regularized decorrelator requires the estimation of noise power. For the case of conventional matched filter, the algorithm exhibits the well-known performance floor due to multiaccess interference. The proposed method with several different front ends are evaluated in Section 5.

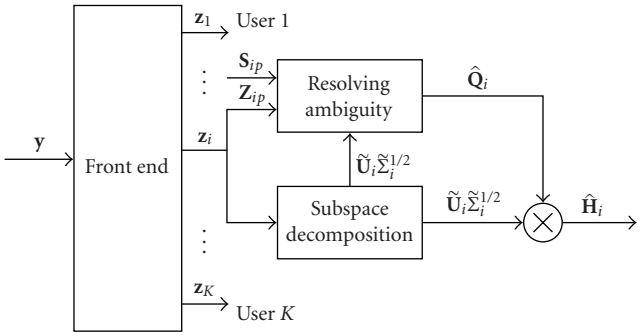


Figure 4: Overall algorithm for blind channel estimation.

The algorithm is derived for the Alamouti coding scheme up to now. However, the proposed method is easily extended to any unitary square block coding that satisfies (11) when the channel length is no less than the codeblock size.

3.5. Computational complexity

The proposed method is described in Figure 4. The main processing consists of the front end, construction and SVD of $\mathbf{R}_i$, and resolving the rotational ambiguity $\mathbf{Q}_i$.

The code matrix in (5) is usually very large for K-user long code CDMA systems. For the case of equal spreading gain G and channel order L between users, the size of $\mathbf{T}$ is approximately $GM \times LMK$, where M is the number of symbols per slot. However, the matrix is very sparse and the number of nonzero elements is approximately $GMLK$ (see Figure 3). The number of operations required for the conventional matched filter front end is given by the number of nonzero elements in $\mathbf{T}$. Hence, the matched filter has approximately $GMLK$ operations. For the decorrelating and regularized decorrelating front end, the inversion of code matrix $\mathbf{T}$ is necessary. Direct inversion is prohibitive for such a large matrix. However, the required inversion can be implemented in an efficient way by utilizing sparsity via the state-space method described in [17]. The computational complexity of the state-space inversion is in the order of GML^2K^2 that is linear with respect to slot size GM in chips.

Since $\mathbf{Z}_{in}$ is an $L \times 2$ matrix and $\mathbf{Z}_{in}\mathbf{Z}_{in}^H$ is Hermitian, the computation of $\mathbf{Z}_{in}\mathbf{Z}_{in}^H$ requires $O(L^2)$ operations. Hence, the construction of $\mathbf{R}_i$ in (19) requires $O(ML^2)$ computations. The SVD of $L \times L$ matrix $\mathbf{R}_i$ can be done with complexity order of L^3. Similarly, the SVD required to resolve the rotational ambiguity has complexity order of constant. Hence, the computational complexity is dominated by the front-end processing and the cost for the required subspace decompositions is negligible.

4. DETECTION

We consider several possible scenarios for symbol detection. First, coherent detection can be done with the estimated channel. We use the output of the front-end processing discussed earlier and perform blockwise maximum likelihood detection to obtain the symbol sequence. Rewriting (8) gives

$$\begin{bmatrix} \mathbf{z}_{i,2n-1} \\ \mathbf{z}_{i,2n}^* \end{bmatrix} = \begin{bmatrix} \mathbf{h}_i^{(1)} & -\mathbf{h}_i^{(2)} \\ \mathbf{h}_i^{(2)*} & \mathbf{h}_i^{(1)*} \end{bmatrix} \begin{bmatrix} s_{i,2n-1} \\ s_{i,2n}^* \end{bmatrix} + \begin{bmatrix} \mathbf{n}_{i,2n-1} \\ \mathbf{n}_{i,2n}^* \end{bmatrix}. \quad (34)$$

Neglecting the color of noise $\mathbf{n}_{i,2n-1}$ and $\mathbf{n}_{i,2n}$, the maximum likelihood estimates for symbol $s_{i,2n-1}$ and $s_{i,2n}$ are given by

$$\begin{bmatrix} \hat{s}_{i,2n-1} \\ \hat{s}_{i,2n}^* \end{bmatrix} = \mathcal{Q}\left(\frac{1}{\beta} \begin{bmatrix} \left(\hat{\mathbf{h}}_i^{(1)}\right)^H & \left(\hat{\mathbf{h}}_i^{(2)}\right)^{\mathrm{T}} \\ -\left(\hat{\mathbf{h}}_i^{(2)}\right)^H & \left(\hat{\mathbf{h}}_i^{(1)}\right)^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathbf{z}_{i,2n-1} \\ \mathbf{z}_{i,2n}^* \end{bmatrix} \right), \quad (35)$$

where $\beta = (\|\mathbf{h}_i^{(1)}\|^2 + \|\mathbf{h}_i^{(2)}\|^2)$ and $\mathcal{Q}$ is the quantization function which selects the symbol vector with minimum distance. Since the covariance of $\mathbf{n}_{i,2n-1}$ and $\mathbf{n}_{i,2n}$ is available, the whitened matched filter detector can be also used instead of (35) for improved performance.

Since the proposed blind method requires only one (space-time) codeblock of pilot symbols for resolving the rotational ambiguity, it is worthwhile to compare its performance with differential demodulation that also requires the same number of pilot symbols. Several authors have proposed noncoherent or differential modulation schemes for space-time coded systems [11, 12]. We consider the differential encoding based on unitary group codes as described in [12]. The encoding procedure is given by the following recursion starting with a (unitary) pilot codeblock $\mathbf{S}_{i1} = \mathbf{S}_{ip}$:

$$\mathbf{S}_{in} = \mathbf{S}_{i,n-1}\mathbf{G}_{in}, \quad (36)$$

where $\mathbf{G}_{in}$ is a unitary matrix belonging to a unitary group $\mathcal{G}$, and carries the information. Although the encoding and decoding steps for the differential scheme are simple for nonspread systems, differential decoding for the CDMA system with multipaths requires additional procedures due to the spreading and intersymbol interference. Similar to [13], we can use a suboptimal two-step approach. First, we apply the front-end processing described in Section 3.1.1 to deal with the despreading and multipath interference, and then use the output of the front end for differential decoding. Since the front-end output (9) has an equivalent signal structure through (nonspread) MIMO channel, we can apply the differential scheme proposed in [12]. Neglecting the color of $\mathbf{N}_{in}$, the detected symbols are given by

$$\hat{\mathbf{G}}_{in} = \arg\max_{\mathbf{G}\in\mathcal{G}} \mathrm{tr}\left(\mathrm{Re}\left(\mathbf{G}\mathbf{Z}_{i,n}^H\mathbf{Z}_{i,n-1}\right)\right). \quad (37)$$

Since front-end processing is the dominant factor in complexity in both cases, the complexity of the coherent and differential schemes is not significantly different for the space-time coded CDMA systems.

5. SIMULATION

In this section, we present some simulation results. First, we evaluate the performance of the proposed channel estimation and detection. For channel estimation, the mean

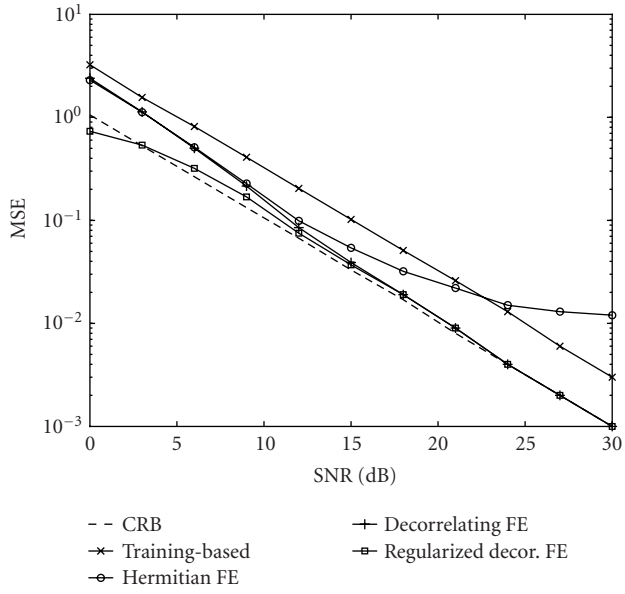


Figure 5: MSE versus SNR; single-user case.

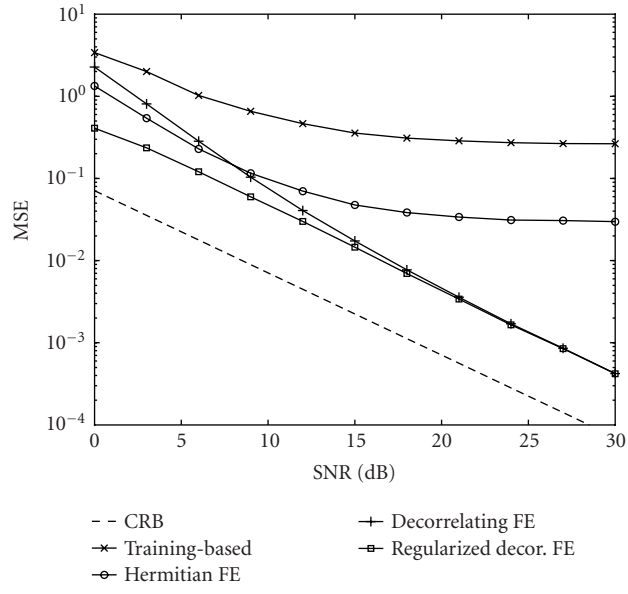


Figure 6: Channel MSE versus SNR; four-synchronous-user case.

square error (MSE) was calculated using Monte Carlo runs and compared with the CRB. For symbol detection, the BER was used. We considered a downlink WCDMA system with two transmit antennas and a single receive antenna. Single ($K = 1$) and multiple BPSK users with equal power were considered. For the multiuser case, we first consider a scenario with ($K = 4$) synchronous users. The spreading codes were randomly generated with spreading gain $G = 32$ and fixed throughout the Monte Carlo simulation for MSE and BER. The slot size $M = 80$ and two pilot symbols, that is, one space-time codeblock, were included at the beginning of the slot of each user. These pilot symbols were used to remove the rotational ambiguity of the blind estimator and to serve as an initial reference in differential detection. For the channel, the block fading model was used, that is, the channel was generated and kept constant over one slot. Since our channel model is deterministic, the channel parameter was fixed during the Monte Carlo runs. For the CRB calculation, the symbol sequence was fixed. For MSE and BER, symbol sequences were generated randomly for each Monte Carlo run. The channel for each TX-RX pair had three fingers $L = 3$. The coefficients are given by $\mathbf{h}^{(1)} = [0.0582 + 0.4331i, 0.1112 + 0.1466i, -0.8375 + 0.2715i]$ and $\mathbf{h}^{(2)} = [0.5317 + 0.1396i, -0.1475 + 0.2831i, 0.6144 - 0.4673i]$. The signal-to-noise ratio (SNR) is defined by $(\|\mathbf{h}^{(1)}\|^2 + \|\mathbf{h}^{(2)}\|^2)GE_c/\sigma^2$, where E_c is the chip energy and σ^2 is the chip noise variance.

We compared the MSE of the proposed channel estimator using different front ends with the CRB and the training-based method. With the availability of the two pilot symbols inserted to resolve the rotational ambiguity, we used the semiblind CRB with a deterministic assumption on data symbols [19]. For the training-based method, a least squares channel estimate was obtained using data corresponding to the pilot symbols. Figure 5 shows the MSE performance for the single-user case. As shown in the figure, the proposed method with the decorrelating and regularized decorrelating front ends closely follows the CRB at high SNR. The proposed method using the conventional matched filter deviates from the CRB as SNR increases due to multipath interference. The least squares estimator based on only pilot symbols is worse than the proposed method with decorrelating or regularized decorrelating front-ends. It does not exhibit a performance floor since it also inverts the submatrix in $\mathbf{T}$ corresponding to the pilot block and eliminates the multipath interference. For the regularized decorrelator, we used the true noise variance and it shows an improved performance at low SNR due to the mitigation of noise enhancement by inversion. Note that the MSE is lower than the CRB. This is because the proposed estimator with the regularized decorrelating front end is not unbiased. Figure 6 shows the MSE for the four synchronous user case where the same channel was used as the single user. In this case, the MSE performance shows a similar behavior with a bigger gap from the CRB. Notice that the absolute value of MSE in this case is smaller than that of the single-user case, whereas the gap between MSE and CRB increases.

We evaluated the BER performance for the coherent detector and the differential scheme in Section 4. For the coherent scheme, we used the whitened version of the ML detector (35). Figure 7 shows the BER performance for the single-user case. For the reference, we used the coherent scheme with the regularized decorrelator and true channel. We observe that the coherent detector with the proposed estimator is marginally better than the differential detector and the difference between different front ends is not significant. Notice that there is about 3 dB SNR loss at BER of 10^{-3} due to channel estimation errors for the coherent detector. Figure 8 shows the BER performance for the four synchronous-user case. The improvement of the proposed method over the

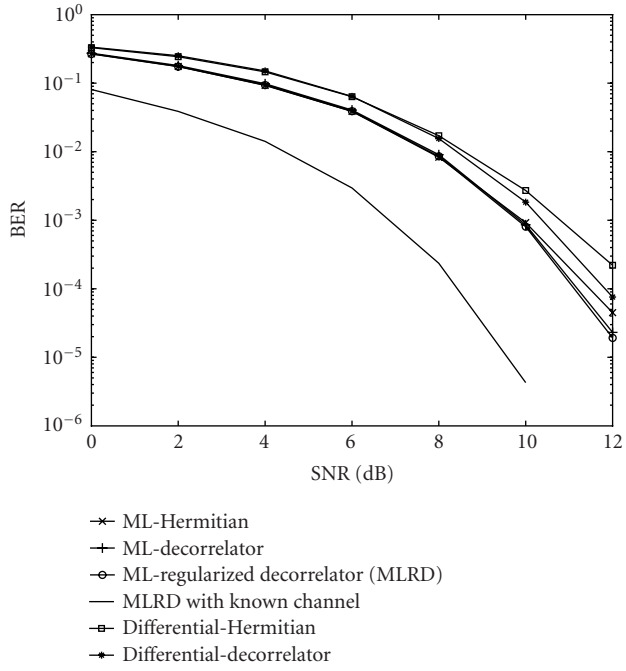


Figure 7: BER versus SNR; single-user case.

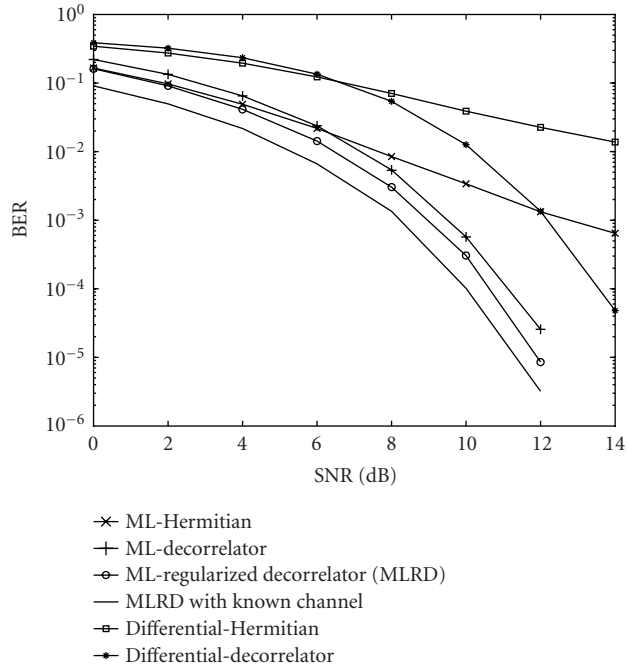


Figure 8: BER versus SNR; four-synchronous-user case.

differential scheme is pronounced. In this case, the difference between perfect channel knowledge and the proposed estimator is less than 1 dB. This is because the proposed method utilizes all user data constructively to estimate the downlink channel, whereas differential detection is performed individually. The performance of the detector using the conventional matched filter becomes worse as SNR increases due to the multiuser interference as expected. As shown in Figure 8, the coherent detection with the proposed channel estimator performs much better than the differential scheme without significant complexity increase or bandwidth efficiency loss when both detectors use the same front end and the same number of pilot symbols for a slot.

Since the proposed algorithm can be used in asynchronous systems without any modification, we evaluated the performance of the proposed method for an asynchronous case. We considered four asynchronous users with long spreading codes. The simulation parameters were the same as in the synchronous case, except that the signals of the users are not synchronized to the slot reference. The delays from the slot reference were 0, 18, 36, 8 chips for the four users. As shown in Figure 9, the performance of the proposed method is almost the same as that in the synchronous case. This is because synchronism between users in the code matrix $\mathbf{T}$ is irrelevant to the front-end processing described in Section 3.1.1. The following subspace technique applies the same to the output of the front end.

Up to now, we considered system parameters that satisfy the identifiability condition well and the proposed method shows a good performance behavior. As discussed in Section 3.2, channel identifiability and the performance of the proposed algorithm depend on the code matrix $\mathbf{T}$. Here, we considered the identifiability condition through simulation. We evaluated the condition number of the code matrix $\mathbf{T}$ as the number of users increases, that is, $\mathbf{T}$ becomes wider. We considered two spreading gains $G = 16, 32$ and different number of users for each spreading gain. The channel length and slot size were fixed as $L = 3$ and $M = 80$. For each pair of spreading gain and number of users, 500 Monte Carlo runs were executed. For each run, the spreading codes were randomly generated for all users, and random delays from the slot reference were generated with the uniform distribution over $[0, G]$ chips independently for each user. Then, matrix $\mathbf{T}$ was formed and the condition number $\kappa(\mathbf{T})$ was calculated. Figure 10 shows the distribution of the calculated condition number of $\mathbf{T}$. The number of outliers ($\kappa(\mathbf{T}) > 200$) were 0, 3, 3, 6 for $K = 2, 3, 4, 5$, with $G = 16$; there was no outlier in any of the cases with $G = 32$. As expected, the condition number for $G = 32$ is smaller than that for $G = 16$, for the same ratio between row and column number of $\mathbf{T}$, since the probability that one spreading code is linearly independent of the others is higher with a larger spreading gain. When the ratio between row and column number approaches one, the condition number suddenly increases. However, for reasonable ratios, the condition number is well distributed with a small mean. This implies that the code matrix $\mathbf{T}$ has full column rank and the proposed method provides good performance for systems with well-designed spreading codes and reasonable loading.

We evaluated the performance of the proposed method when the system is heavily loaded. We considered the number of users $K = 8, 10$ (each user had a randomly generated

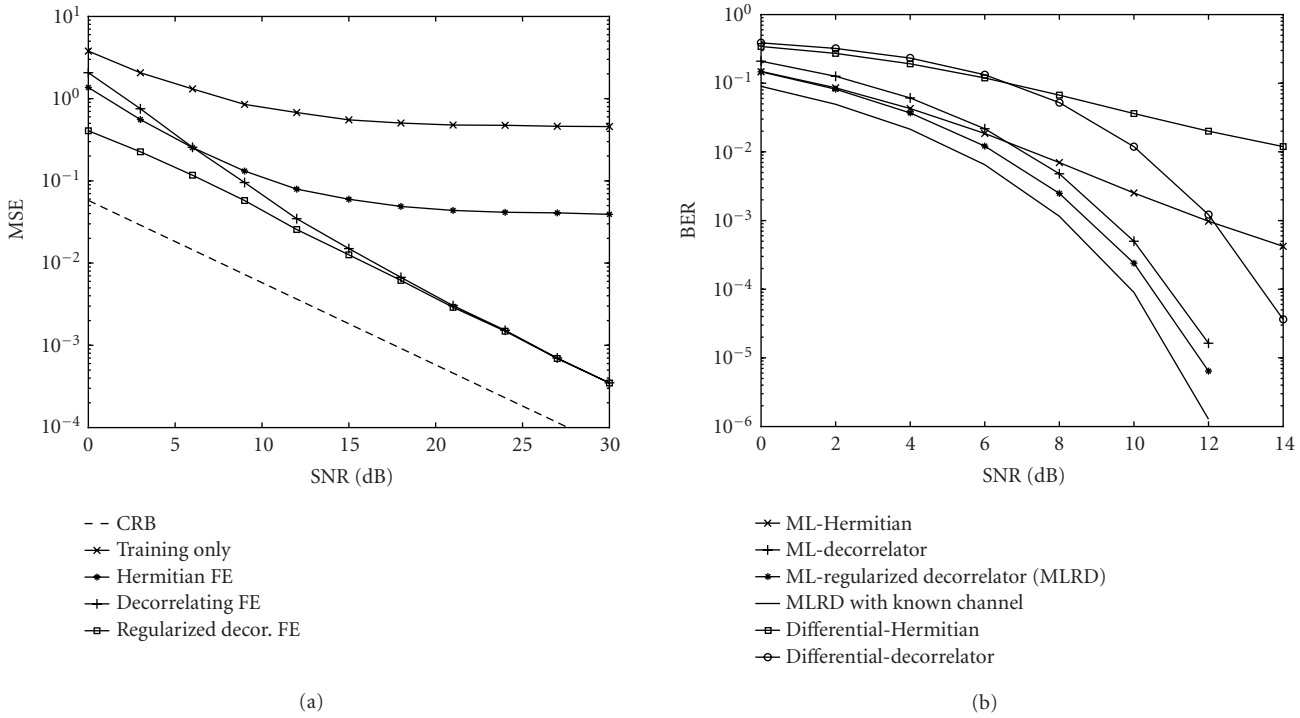


FIGURE 9: Four-asynchronous-user case ($G = 32$, $M = 80$, $L = 3$, $D = [0, 18, 36, 8]$). (a) MSE versus SNR and (b) BER versus SNR.

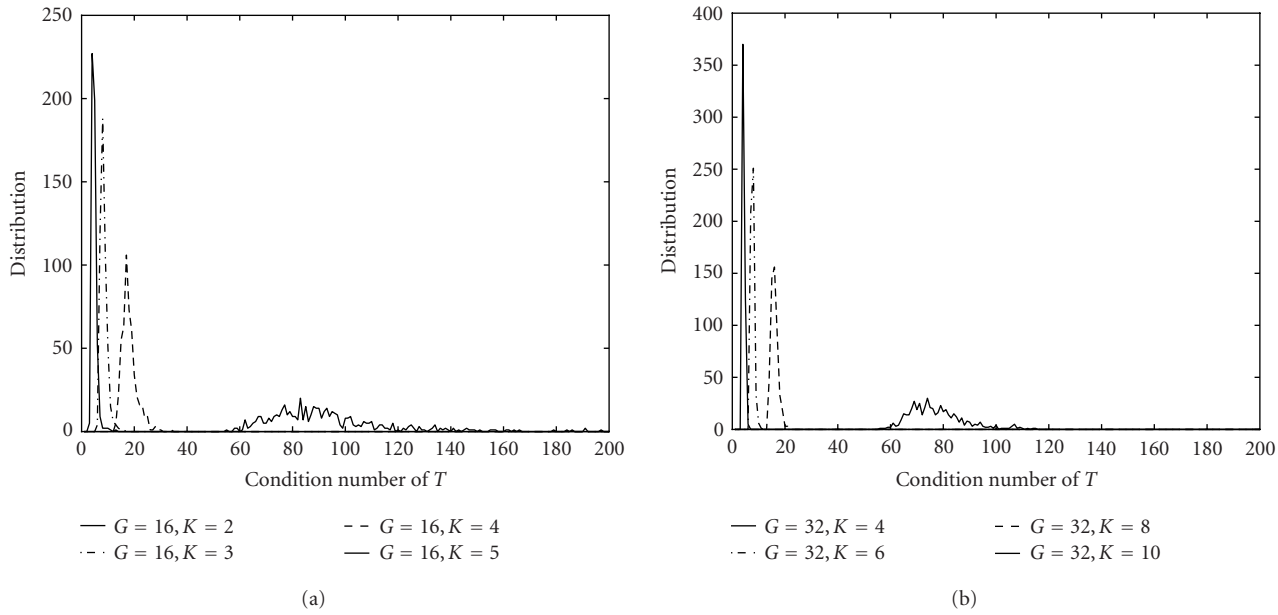


FIGURE 10: Distribution of the condition number of $\mathbf{T}$ ($M = 80$, $L = 3$). (a) $G = 16$ and (b) $G = 32$.

spreading code); all other simulation parameters were the same as in Figure 8. (In this cases, the code matrix $\mathbf{T}$ is almost square but still tall.) Figure 11 shows the BER performance of the coherent detector with the proposed estimate and the differential detector. Performance degrades as the number of users increases. In particular, the performance with the decorrelating front end deviates much from that of the regularized decorrelator due to noise enhancement by

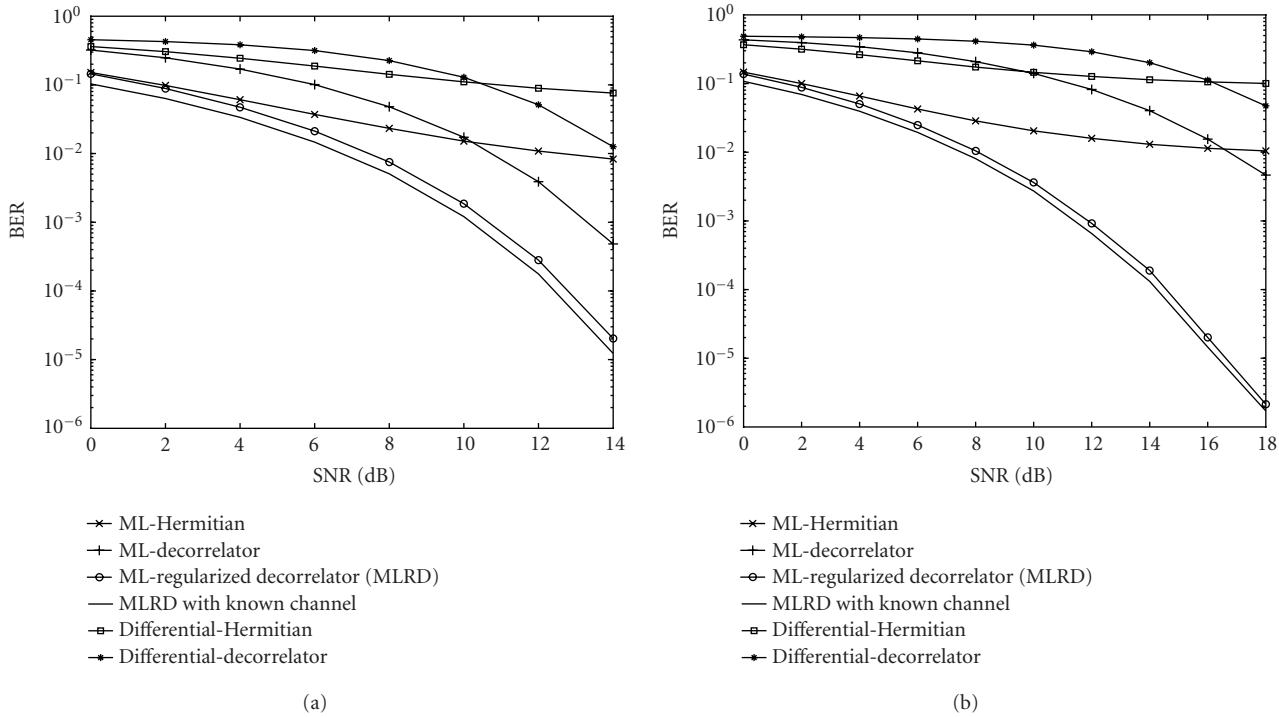


Figure 11: BER versus SNR; heavily loaded cases ($G = 32$, $M = 80$, $L = 3$). (a) $K = 8$ and (b) $K = 10$.

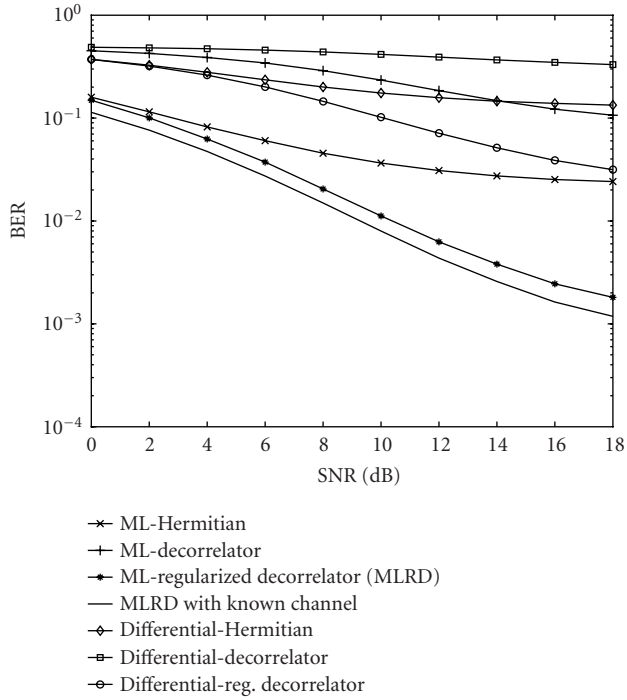


Figure 12: BER versus SNR; overloaded case ($G = 32$, $M = 80$, $L = 3$, $K = 12$).

the inversion. (See Figure 10b for the condition number of $\mathbf{T}$.) In the case of the regularized decorrelator, the noise enhancement is mitigated and the coherent detector using the proposed channel estimate has almost the same performance as the detector with known channel. Finally, we considered an overloaded case. The number of users was chosen to be $K = 12$ with spreading gain $G = 32$ and channel length $L = 3$. ($G < LK$.) Since the proposed method can accommodate $K \leq \lfloor G/L \rfloor$ users in the code matrix $\mathbf{T}$, the additional two users are modeled as additive noise. Figure 12 shows the average BER of the ten users modeled in the code matrix. As expected, the proposed method shows a performance floor due to multiuser interference as the SNR increases. The performance floor can be lowered by using larger spreading gain. For the same front end, however, the coherent detector with the proposed channel estimate shows better performance than the differential scheme.

6. CONCLUSION

We proposed a new blind channel estimation technique for space-time coded CDMA systems. A new identifiability condition was established. The proposed method identifies the channel of each transmit-receive pair simultaneously, exploiting the subspace structure of CDMA signals and the orthogonality of space-time codes; it requires only few pilot symbols. The performance of the proposed method is evaluated through simulation and is compared with that of differential schemes. The proposed algorithm can be also applied to general unitary space-time coding schemes.

APPENDIX

PROOF OF PROPOSITION 2

The proof is the complex-valued version of the one in [18]. Let $\mathbf{A} \triangleq \mathbf{Z}_{ip}\mathbf{S}_{ip}^H$ and $\mathbf{B} \triangleq (\alpha_{i1}/\sqrt{\alpha_i})\widetilde{\mathbf{U}}_i\widetilde{\Sigma}_i^{1/2}$. Then, (24) is written as

$$\left\|\mathbf{Z}_{ip}\mathbf{S}_{ip}^H - \frac{\alpha_{i1}}{\sqrt{\alpha_i}}\widetilde{\mathbf{U}}_i\widetilde{\Sigma}_i^{1/2}\mathbf{Q}\right\|_F^2 = \|\mathbf{A}-\mathbf{B}\mathbf{Q}\|_F^2,$$

$$\begin{aligned}\|\mathbf{A}-\mathbf{B}\mathbf{Q}\|_F^2 &= \operatorname{tr}\left((\mathbf{A}-\mathbf{B}\mathbf{Q})^H(\mathbf{A}-\mathbf{B}\mathbf{Q})\right)\\ &= \operatorname{tr}\left(\mathbf{A}^H\mathbf{A}+\mathbf{B}^H\mathbf{B}-\mathbf{Q}^H\mathbf{B}^H\mathbf{A}-\mathbf{A}^H\mathbf{B}\mathbf{Q}\right)\\ &= \operatorname{tr}\left(\mathbf{A}^H\mathbf{A}+\mathbf{B}^H\mathbf{B}\right)-2\operatorname{tr}\left(\operatorname{Re}\left(\mathbf{Q}^H\mathbf{B}^H\mathbf{A}\right)\right).\end{aligned} \tag{A.1}$$

Since $\mathbf{A}$ and $\mathbf{B}$ are given, the optimization problem is equivalent to maximizing $\operatorname{tr}(\operatorname{Re}(\mathbf{Q}^H\mathbf{B}^H\mathbf{A}))$. Let the SVD of $\mathbf{B}^H\mathbf{A}$ be given by

$$\mathbf{B}^H\mathbf{A} = \mathbf{U}_Q\Sigma_Q\mathbf{V}_Q^H. \tag{A.2}$$

Then, we have

$$\begin{aligned}\operatorname{tr}\left(\operatorname{Re}\left(\mathbf{Q}^H\mathbf{B}^H\mathbf{A}\right)\right) &= \operatorname{tr}\left(\operatorname{Re}\left(\mathbf{Q}^H\mathbf{U}_Q\Sigma_Q\mathbf{V}_Q^H\right)\right)\\ &= \operatorname{tr}\left(\operatorname{Re}\left(\mathbf{V}_Q^H\mathbf{Q}^H\mathbf{U}_Q\Sigma_Q\right)\right)\\ &= \operatorname{tr}\left(\operatorname{Re}\left(\mathbf{X}\Sigma_Q\right)\right)\\ &= \operatorname{Re}\sum_{j=1}^{2}x_{jj}\sigma_j,\end{aligned} \tag{A.3}$$

where $\Sigma_Q = \operatorname{diag}(\sigma_1,\sigma_2)$ and $\mathbf{X} \triangleq \mathbf{V}_Q^H\mathbf{Q}^H\mathbf{U}_Q$. Since $\mathbf{Q}$ is unitary, $\mathbf{X}$ is also unitary. Hence, we have $|x_{jj}| \le 1$. The maximum of $\operatorname{tr}(\operatorname{Re}(\mathbf{Q}^H\mathbf{B}^H\mathbf{A}))$ occurs when $x_{jj} = 1$ since $\sigma_j \ge 0$ for all j. This implies that $\mathbf{X}$ is an identity matrix, which concludes the proof.

ACKNOWLEDGMENTS

This work is supported in part by the Multidisciplinary University Research Initiative (MURI) under the Office of Naval Research Contract N00014-00-1-0564. It was prepared through collaborative participation in the Communications and Networks Consortium sponsored by the US Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011. The US Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon.

REFERENCES

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.

[2] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. URSI International Symposium on Signals, Systems, and Electronics (ISSSE '98)*, pp. 295–300, Pisa, Italy, September 1998.

[3] H. Huang, H. Viswanathan, and G. J. Foschini, "Achieving high data rates in CDMA systems using BLAST techniques," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM '99)*, vol. 5, pp. 2316–2320, Rio de Janeiro, Brazil, December 1999.

[4] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–596, 1999.

[5] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.

[6] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, 1999.

[7] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Transactions on Information Theory*, vol. 48, no. 7, pp. 1804–1824, 2002.

[8] P. Stoica and G. Ganesan, "Space-time block codes: trained, blind and semi-blind detection," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP '02)*, vol. 2, pp. 1609–1612, Orlando, Fla, USA, May 2002.

[9] A. L. Swindlehurst and G. Leus, "Blind and semi-blind equalization for generalized space-time block codes," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2489–2498, 2002.

[10] N. Ammar and Z. Ding, "On blind channel identifiability under space-time coded transmission," in *Proc. 36th IEEE Annual Asilomar Conference on Signals, Systems, and Computers*, vol. 1, pp. 664–668, Pacific Grove, Calif, USA, November 2002.

[11] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 7, pp. 1169–1174, 2000.

[12] B. L. Hughes, "Differential space-time modulation," *IEEE Transactions on Information Theory*, vol. 46, no. 7, pp. 2567–2578, 2000.

[13] H. Li and J. Li, "Differential and coherent decorrelating multiuser receivers for space-time-coded CDMA systems," *IEEE Trans. Signal Processing*, vol. 50, no. 10, pp. 2529–2537, 2002.

[14] K. Rohani and L. Jalloul, "Orthogonal transmit diversity for direct spread CDMA," in *Proc. ETSI SMG2*, Stockholm, Sweden, September 1997.

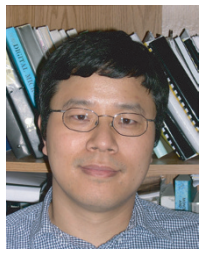
[15] B. Hochwald, T. L. Marzetta, and C. B. Papadias, "A transmitter diversity scheme for wideband CDMA systems based on space-time spreading," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 1, pp. 48–60, 2001.

[16] H. Li, X. Lu, and G. B. Giannakis, "Capon multiuser receiver for CDMA systems with space-time coding," *IEEE Trans. Signal Processing*, vol. 50, no. 5, pp. 1193–1204, 2002.

[17] L. Tong, A.-J. van der Veen, P. Dewilde, and Y. Sung, "Blind decorrelating RAKE receivers for long-code WCDMA," *IEEE Trans. Signal Processing*, vol. 51, no. 6, pp. 1642–1655, 2003.

[18] G. H. Golub and C. F. van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, Md, USA, 1989.

[19] E. De Carvalho and D. T. M. Slock, "Cramer-Rao bounds for semi-blind, blind and training sequence based channel estimation," in *Proc. 1st IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications (SPAWC '97)*, pp. 129–132, Paris, France, April 1997.



Youngchul Sung received the B.S. and M.S. degrees from Seoul National University, Seoul, Korea, in electronics engineering in 1993 and 1995, respectively. He is currently a Ph.D. candidate at Adaptive Communications and Signal Processing Laboratory, School of Electrical and Computer Engineering, Cornell University, Ithaca, New York. He worked as a Research Engineer at Goldstar Information and Communications, Ltd., Seoul, Korea, from 1995 to 2000. At Goldstar, he developed high-data-rate wireless transmission systems and wireless local loop based on wideband code-division multiple-access (CDMA). His areas of interest include statistical signal processing and communication networks. He is currently interested in signal processing techniques for large-scale sensor networks.

Lang Tong received the B.E. degree from Tsinghua University, Beijing, China, in 1985, and the M.S. and Ph.D. degrees in electrical engineering in 1987 and 1990, respectively, from the University of Notre Dame, Notre Dame, Indiana. He was a Postdoctoral Research Affiliate at the Information Systems Laboratory, Stanford University, in 1991. Currently, he is an Associate Professor in the School of Electrical and Computer Engineering, Cornell University, Ithaca, New York. Dr. Tong received the Young Investigator Award from the Office of Naval Research in 1996 and the Outstanding Young Author Award from the IEEE Circuits and Systems Society. His areas of interest include statistical signal processing, adaptive receiver design for communication systems, signal processing for communication networks, and information theory.

Ananthram Swami received the B.Tech. degree from IIT, Bombay, the M.S. degree from Rice University, and the Ph.D. degree from the University of Southern California (USC), all in electrical engineering. He has held positions in Unocal Corporation, USC, CS-3, and Malgudi Systems. He is currently with the US Army Research Laboratory where he is a Fellow. His work is in the broad area of signal processing and communications. He is a Member of the IEEE Signal Processing (SP) Society's Technical Committee on SP for Communications, an Associate Editor of the IEEE Transactions on Wireless Communications and of the IEEE Transactions on Signal Processing, and a Member of the Editorial Board of the IEEE SP Magazine (SPM). He has served as an Associate Editor for IEEE SP Letters and IEEE Transactions on Circuits & Systems-II. He was the Coorganizer and Cochair of the 1993 IEEE-SPS HOS Workshop, the 1996 IEEE-SPS SSAP Workshop, and the 1999 ASA-IMA Workshop on Heavy-Tailed Phenomena. He is a Coguest Editor of a 2004 special issue of IEEE SPM on "signal processing for networking." He was a Statistical Consultant to the California Lottery, developed a Matlab-based toolbox for non-Gaussian signal processing, and has held visiting faculty positions at INP, Toulouse, France.